DOCUMENT RESUME

ED 202 402 HE 013 923

AUTHOR	Williams, Albert P.
TITLE	Predicting Performance in Medical Education Continuum: Toward Better Use of Conventional Measures.
INSTITUTION	Rand Corp., Santa Monica, Calif.
SPONS AGENCY	Health Resources Administration (DHEW/PHS), Bethesda, Md.; National Institutes of Health (DHEW), Bethesda, Md.
PUB DATE	Apr 80
CONTRACT	HRA-231-75-0614; N01-MB-24196; NIH-72-4196
NOTE	43p.
AVAILABLE FROM	The Rand Corporation, 1700 Main Street, Santa Monica, CA 90406 ($3.00).
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	*Admission Criteria; Affirmative Action; Aptitude Tests; Certification; *College Entrance Examinations; Comparative Analysis; Decision Making; *Grade Point Average; Higher Education; In State Students; Medical Education; *Medical Schools; *Minority Groups; Out of State Students; Place of Residence; Predictive Measurement; *Predictor Variables; Science Tests; Selective Admission; Whites
IDENTIFIERS	Creighton University NE; *Duke University NC; Medical College of Wisconsin; Ohio State University; Tufts University MA; University of California Los Angeles; University of California San Francisco; University of Florida; University of New Mexico; University of Vermont

ABSTRACT

Medical school admissions and performance in 10 medical schools were assessed in relation to prediction using conventional measures. The origin of the research was an attempt to determine the effects of affirmative action on academic medicine. For the 10 schools, admissions decisions were analyzed, and an attempt was made to model statistically what factors were used in the decision-making. Additional areas of analysis were the effect of the state of residence on the chance of an individual's being admitted to at least one school, the performance of minority and majority group students on the National Boards, Part I and II, and performance on specialty board exams, specifically the exam of the American Board of Internal Medicine (ABIM). It was found that only three schools fail to use the science score on the Medical College Admission Test (MCAT) for admitting minorities as well as the majority. Data are presented that show admission probabilities in 1975 for two hypothetical candidates (minority and majority group members). The admissions probabilities were found to vary by race. It was also found that a student's place of residence strongly affected the probabilities of being admitted, and that some of the schools were weighting the state of residence more heavily than they realized. The assessment of performance on the National Boards revealed that for majority students the most significant predictor was science grade point average (GPA), followed by science MCAT, followed by selectivity of an undergraduate school, followed by general information MCAT, verbal MCAT, and quantitative MCAT. For minority students the verbal MCAT was better than the science MCAT or science GPA. For the ABIM, the effects of MCAT, undergraduate GPA, and National Boards were significant even after controlling for rank and class, age, and school. (SW)

ED202402

HE

PREDICTING PERFORMANCE IN MEDICAL EDUCATION CONTINUUM: TOWARD BETTER USE OF CONVENTIONAL MEASURES

Albert P. Williams

April 1980

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Paul Cooper

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

HE 013 923

P-6555

The Rand Paper Series

Papers are issued by The Rand Corporation as a service to its professional staff. Their purpose is to facilitate the exchange of ideas among those who share the author's research interests; Papers are not reports prepared in fulfillment of Rand's contracts or grants. Views expressed in a Paper are the author's own, and are not necessarily shared by Rand or its research sponsors.

The Rand Corporation
Santa Monica, California 90406

Predicting Performance in Medical Education Continuum: Toward Better Use of Conventional Measures [1]

The origin of this research[2] was a study for the federal government. It addressed a very broad question: What are the effects of federal programs on academic medicine. We enlisted--conned, perhaps--ten medical schools into participating in this study. Duke was one of them. The other nine were UCLA, UCSF, Ohio State, Medical College of Wisconsin, New Mexico, Florida, Vermont, Tufts, and Creighton. Without defending that sample, I'll say that we at least crunched some numbers to make us think we covered a broad spectrum of schools that differed on a number of dimensions.

Before turning to our analysis of medical school admissions and performance, I should emphasize--if that isn't clear--that I don't have any norms regarding what ought to be the factors used in admissions, or, indeed, how performance ought to be measured. What my colleagues and I are trying to do in this analysis is to describe statistically what is happening. [3]

---

[1] Presented on 25 April 1980 at a Conference on Premedical Preparation and Medical School Admissions sponsored by the Josiah Macy, Jr. Foundation at the Duke University Medical Center, Durham, North Carolina.

[2] Research described in this talk was performed under three contracts to the Department of Health, Education, and Welfare: NIH 72-4196; NO1-MB-24196; HRA-231-75-0614.

[3] R-1464-HEW, Federal Manpower Legislation and the Academic Health Centers: An Interim Report, G. M. Carter, D. S. Chu, J. E. Koehler, R. L. Slighton, A. P. Williams, April 1974; R-2014-HEW, The Effect of State of Residence on Medical School Admissions, J. E. Rolph, A. P. Williams, C. Lee, February 1978; R-2029-HEW, Predicting Minority and Majority Medical Student Performance on the National Board Exams, J. E. Rolph, A. P. Williams, A. L. Laniear, W. D. Cooper, November 1978; R-2030-HEW, Fac-

$3.00

Now, if Bob Beran was bringing you up to date, I will be bringing you back to date, because the data I'm going to talk about are from the old MCAT. For some fairly obvious reasons, the studies have only begun with the new MCAT. If you're looking very far down the pike, you can't do so with the new MCAT. I think it'll be fascinating to do that when the data are available.

Standardized tests clearly have an effect on the people who go through the medical education continuum. (First Slide) They serve as toll gates or, if you prefer, hurdles. The next proposition might be a little bit more controversial. I believe the results of our analysis suggest that tests are of more than ritual value. There's no question about their ritual value. They determine how people progress. In my view, the studies that have failed to find any relationship between these tests and other evaluation criteria usually fail to find anything because they've used small numbers and, in many cases, weak methods. It's very hard with the small numbers that you get from a class in an individual school to do an analysis that teases out, if you will, significant relationships. Before proceeding further, I want to be clear that, in talking about quantitative measures, I do not want to deemphasize the importance of qualitative data. I think much more research needs to be done on qualitative performance measures.

---

tors Affecting Medical School Admission Decisions for Minority and Majority Applicants: A Comparative Study of Ten Schools, A. P. Williams, W. D. Cooper, C. L. Lee, December 1979

PROPOSITIONS ABOUT VALUE OF TESTS IN MEDICAL EDUCATION

- Standardized tests (MCAT, National Boards, Certifying Exams) clearly affect entry to and progress through medical education continuum

- Results of these tests are of more than ritual value

- Studies using test data have tended to use small numbers and weak methods

- More research on qualitative attributes is needed

I'm going to try to talk briefly, in summary, about four different studies (Slide 2) as examples of analysis that we have done. First, for all these ten schools, we analyzed admissions decisions. That is to say, we tried to model statistically what factors they used in determining who gets in and who doesn't. Second, we looked nationally at data from all schools to assess the effect of the state of residence on the chance of an individual's being admitted to at least one school. The criterion there is "I'll get in some place." Third, we analyzed minority and majority student performance on the National Boards, Part I and II. Finally, for a set of classes in our nine schools, we analyzed performance on specialty board exams, specifically the exam of the American Board of Internal Medicine, which is an exam given prior to certification in medicine.

Now, let me go back to the admissions process, because that's what I will talk about first. As I suggested, we talked with admissions deans; we sat in on committee meetings; etc., to try to get the feel of how this process worked. The data we used for analysis were data from all ten schools from the American Medical College Application Service (AMCAS). Our data set includes what the admissions dean or admissions committee has for the objective, uniform data at all schools.

We take data on all applicants to each of those schools, and we estimate a model for each school that predicts the probability of admission as a function of individual characteristics. The next slide shows the significance of various factors. (Slide 3) I will only list the most important ones. We estimated two

EXAMPLES OF STATISTICAL ANALYSIS USING CONVENTIONAL PERFORMANCE MEASURES

- Individual school admissions decisions
- Effect of state of residence on admission chances
- Minority and majority student performance on National Boards
- Performance on specialty board exams

models: one for the minority applicants, and another for the majority or white applicants. Rather than giving you a lot of numbers, I use a simple color code: red shows it is a significant factor only in predicting minority admission; blue shows it is significant for the majority; and yellow, if significant, for both. Not surprisingly, you see that science GPA is yellow, hence significant for all schools for both minority and majority applicants. When you go down to nonscience GPA you see that the picture there is much less uniform. The row for science hours addresses one of the questions that came up earlier this morning: What is the effect of the number of science courses you take on your probability of admission? The answer is that, controlling for these other things, there is little effect.

Let me note that this logit analysis is similar to multiple regression. These are not simple correlations of one item with admissions; they are controlling for other things. For example, science hours, by themselves, are surely correlated with chances of admission to medical school. But after you control for how well people do on the science MCAT, then science hours wash out, if you will, and the performance on the MCAT or the performance on the GPA are the true predictors.

By going down the figure, you can see, for example, that verbal MCAT is used by only one school for admitting minorities; that is to say, it is only significant ($p < .1$) for one school. The same is true for quantitative MCAT. No school used this for admitting minorities; no school used the general information MCAT

FACTORS IN ADMISSIONS

VARIABLE	SCHOOL 1	2	3	4	5	6	7	8	9	10
SCIENCE GPA	Y	Y	Y	Y	Y	Y	Y	Y	Y	B
NON-SCIENCE GPA	B		B		B	Y	B	R		R
SCIENCE HOURS*		R				B		B		
VERBAL MCAT	B	B	B	Y	B	B	B	B	B	
QUANTITATIVE MCAT	B	B			Y	B	B	B		
GENERAL INFO MCAT				B	B	B		B		
SCIENCE MCAT	B	Y	Y	R	Y	Y	Y	Y	B	Y
SAME UNDERGRADUATE	B	B	B	Y	B	B	Y	Y	Y	
SELECTIVITY		B	Y		Y		Y	Y	Y	
STATE RESIDENT	Y	Y	Y	B	B	Y	Y	B	B	Y

R = SIGNIFICANT FACTOR ONLY FOR MINORITY APPLICANTS

B = SIGNIFICANT FACTOR ONLY FOR MAJORITY APPLICANTS

Y = SIGNIFICANT FACTOR FOR BOTH MINORITY AND MAJORITY APPLICANTS

= NOT SIGNIFICANT FOR EITHER AT 0.1 LEVEL

SLIDE 3

for admitting minorities; and so on. However, science MCAT is very widely used. Only three schools fail to use the science MCAT for admitting minorities as well as the majority. School four used science MCAT for minorities but not majorities. All the rest used it either for both or for one.

The next line is the one that shows the effects of being from the same undergraduate school. There is strong preference for graduates of the same undergraduate school.

Let me briefly explain selectivity. That is based on Barron's Profile of American Colleges and Universities, with which you may be familiar. Basically, it is an independent rating of the selectivity of the undergraduate school. We put this in the equation to see whether or not one is favored more for having gone to more academically prestigious and selective undergraduate schools.

The bottom line is home state residence. Note that being a home state resident is very important for majorities for all medical schools, irrespective of whether or not it is a public or a private school. Some of the schools, however, do not attach significance to state of residence in the case of minority applicants.

That is a summary of the most significant factors. We had other variables in the model, but those are the ones that are the most significant and apply across all schools. In effect, this analysis shows the implicit weighting that the medical schools attach to these quantitative or quantifiable variables in the admissions process.

Now I want to make a couple of things clear, and I think it may speak to, if you will, the over emphasis or under emphasis on quantifiable factors. The model we use here is one that I think uses as sophisticated statistics as the data will permit. The best we can do is still not very good at predicting admissions. Let me be clearer. Think of the multiple regression analog, R-squared. The amount of variance explained by the quantifiable or quantitative factors is never more than 60%. So there is something else going on in the evaluation of students that is subjective. It deals with factors that are, if not noncognitive, at least not measurable, by the routine measures. In short, admissions committees take into account more than just the numbers.

Now, our analysis shows another thing. Medical schools are not always taking into account what they think they are. I have briefed the deans of admissions at each of these ten schools. They are sometimes surprised, particularly at the importance they may attach to a variable such as state of residence. They often are not really aware of the implications of setting aside, say, 10 or 20 or 30 places for home state residents; they don't know what they are trading off there.

Now, let me go to some analysis of hypothetical applicants. The original purpose of this analysis was to look at affirmative action. As I have shown, there are two models--one for minority and one for majority admissions. When the federal government funded us, they wanted to know what medical schools were doing by way of affirmative action. Therefore, we used these models to

make some comparisons of the probabilities of admission by race. Let me explain how this is done. (Slide 4) This table shows admission probabilities in 1975 for two hypothetical candidates to a private school from our sample. I have called these for short hand, "good minority" and "good majority" applicants. "Good," in this case, means someone whose characteristics on all the quantitative factors are one-half standard deviation above the mean of the group of minority or majority applicants, as the case may be, to that school in 1975. Using the school's admissions equation, we calculate the estimated probability of that hypothetical applicant's being admitted. The minority applicant has roughly a .55 chance of being admitted, if judged as a minority applicant. The majority applicant has about a .04 chance of being admitted, if judged by the majority process. To examine the effect of affirmative action admissions, we assess what might happen if the school did not know that the hypothetical minority applicant had minority status. The probability falls from .55 to .001. If a hypothetical majority applicant indicated that he was a minority applicant and were evaluated by affirmative action standards, his probability would go from .04 to .95. This seemed to us a useful measure of the effect of affirmative action programs for the schools and for the government. (Slide 5) The next table shows a similar analysis for a public school in our group. However, in the previous table, the applicant was an out-of-state applicant; this is a home-state applicant.

ADMISSION PROBABILITIES VARY BY RACE

1975 OUT OF STATE PRIVATE SCHOOL CANDIDATES

	MINORITY PROCESS	MAJORITY PROCESS
GOOD MINORITY APPLICANT	0.547	0.001
GOOD MAJORITY APPLICANT	0.941	0.038

ADMISSION PROBABILITIES VARY BY RACE

1975 PUBLIC SCHOOL HOME STATE RESIDENT

	MINORITY PROCESS	MAJORITY PROCESS
GOOD MINORITY APPLICANT	0.387	0.004
GOOD MAJORITY APPLICANT	0.668	0.121

The next thing I would like to talk about is our analysis of the effect of state of residence on the probability of admission. (Slide 6) The outcome variable is the applicant's success in getting into any school. So if I applied at ten schools and I get into one, the dependent variable for me takes a value of one; and if I get admitted nowhere, it is zero. If I get into four schools, it still takes the value of one.

This analysis is similar to that shown in the multicolored table that I showed before. Our analysis uses MCATs, GPAs, and other characteristics for each applicant. But now we use them for all applicants to all schools, so we have this great humongous discriminate analysis of 35,000 to 37,000 applicants. However, we are now also including another variable that says that Al Williams is an applicant from California, or Al Williams is an applicant from South Dakota. We estimate the effect of being from that state, controlling all of my other characteristics. Because we observed in our earlier analysis that the minority and majority processes were different, we estimate separate minority and majority equations. (Slide 7) This table shows the effect of state of residence on admission probabilities for hypothetical applicants from sample states. We see, for example, that the average applicant from South Dakota had a .81 probability of being admitted to medical school. Now by contrast that applicant from California had about a .13 chance of being admitted. Now, if we jack up the characteristics to one-half a standard deviation above the mean--our "good" applicant from before--then the difference narrows as you might

SLIDE 6

THE INDIVIDUAL APPLICANT—A NATIONAL ANALYSIS OF 1972-1975

- Outcome measure is admission to at least one medical school
- Similar applicant characteristics as in the ten school analysis are used to predict outcome
- The minority and majority admission processes are analyzed separately
- The emphasis is on understanding the state of residence effect on the admission odds

SLIDE 7

SELECTED STATE OF RESIDENCE EFFECTS
(Majority Applicants 1975)

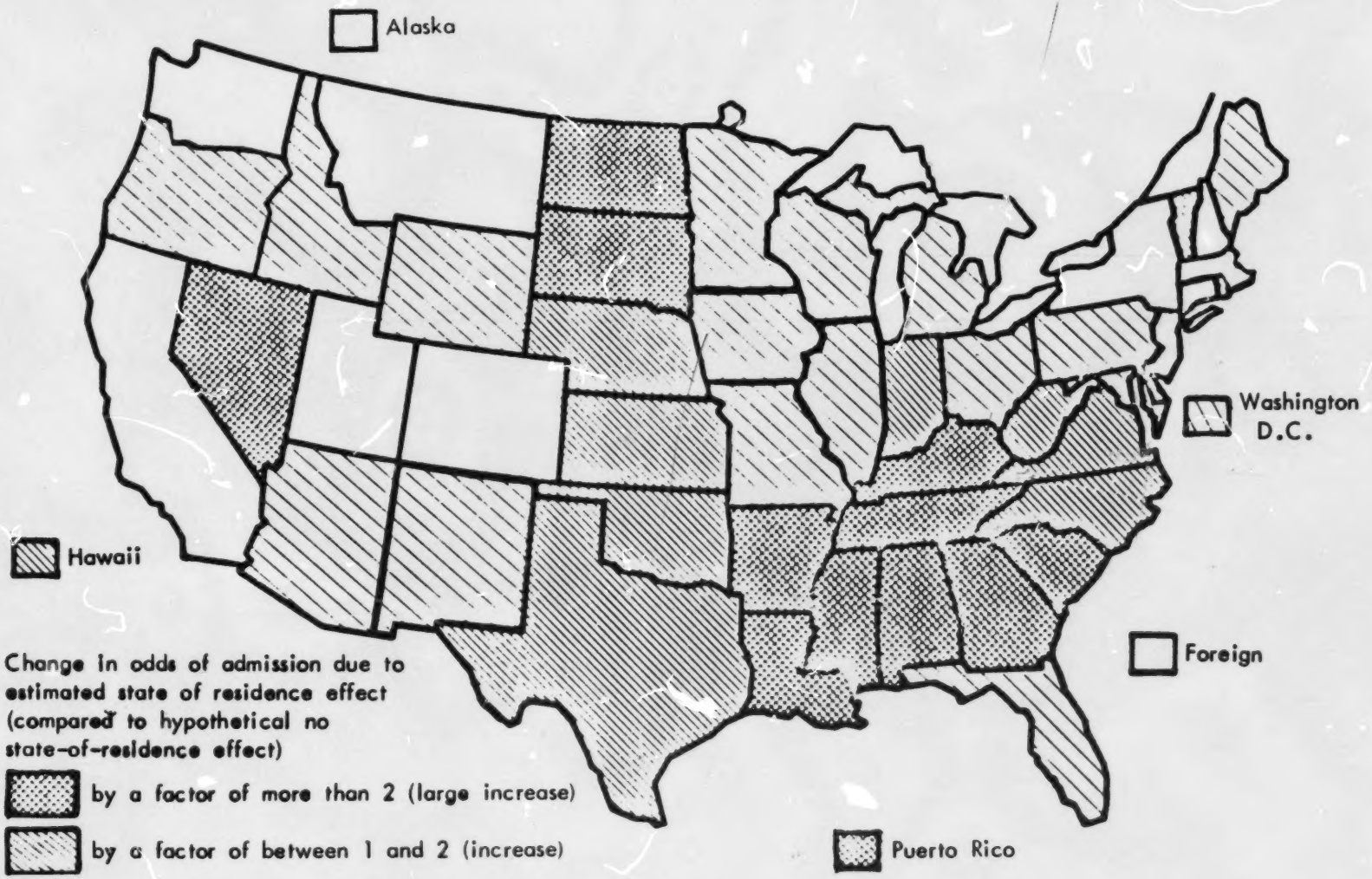
State (Rank)	Multiplicative Odds Factor	Average Probability of Admission	Probability of Admission for a "Good Applicant"
South Dakota (1)	7.68	0.812	0.949
Virginia (16)	1.50	0.456	0.784
Florida (36)	0.68	0.277	0.623
California (52)	0.27	0.129	0.391

expect. The Californian has about a .39 probability of being admitted, and the South Dakotan has about a .95 probability. The bottom line is that it makes one heck of a lot of difference where you come from. Remember we are controlling for all other characteristics.

Now let me just quickly put up a map which shows this effect by state. (Slide 8) The lighter the state, the lower the chances that an applicant with a given characteristic will be admitted to some school. Now this map is based on how residence affects your odds of being admitted. If the odds are multiplied by a factor of .5 and if your odds of being admitted are even (1 to 1) with no state-of-residence effect, then they go down to at least one out of two, and so on. The darkest area here shows states from which applicant's odds are increased by a factor of two or more.

It makes a lot of difference where you come from on your chance of being admitted to medical school. I think this was a finding that I think surprised a lot of people. Everyone knew that each school discriminated against people from out of state, but many felt that such discrimination more or less all balanced out. "They discriminate against our applicants and we discriminate against theirs." It doesn't balance out at all; it makes a lot of difference what your home state is.

Now what factors account for this? (Slide 9) There are two: <u>first</u>, and not surprising, is the number of medical school places in the state. Second, and perhaps less obvious, is the physician/patient ratio in the state.


Alaska
Washington D.C.
Hawaii
Foreign
Puerto Rico
Change in odds of admission due to estimated state of residence effect (compared to hypothetical no state-of-residence effect)
by a factor of more than 2 (large increase)
by a factor of between 1 and 2 (increase)
by a factor of between 1 and .5 (decrease)
by a factor of less than .5 (large decrease)

Effects of state of residence on a majority applicant's odds of admission to medical school (average 1972 - 1975)

FACTORS IN STATE OF RESIDENCE EFFECTS

- MORE MEDICAL SCHOOL PLACES PER CAPITA
- FEWER PHYSICIANS PER CAPITA

Now, let me talk about performance analysis that Syd mentioned. (Slide 10) The dependent variables in our analysis are scores from Parts I and II of the National Board of Medical Examiners exam. I think most of you are familiar with these. Part I--the basic science group of tests--includes separate tests for anatomy, behavioral science, biochemistry, microbiology, pathology, physiology, and pharmacology. Part II has comparable separate tests for the clinical curriculum-- medicine, surgery, and so forth. We are predicting national board scores of students using their preadmission data and data on the medical school they attend. Now we have nine schools--not ten as the slide indicates. Because we were looking at minority and majority performance and one of the original ten did not have a minority student who graduated during that period (though they had some enrolled), we dropped out both classes from that school for those years. Our analysis deals with about 1900 majority students and about 200 minority students.

The variables used for predicting performance are the same ones that were used in admissions because the basic purpose of the analysis is to see how well the preadmissions characteristics predict performance. Now we also used a variable to take account of each medical school's effect on the performance on the test. So, for example, there is a Duke variable. If the sign on the Duke variable were positive and the coefficient were 50 for the biochemistry equation, that Duke student, other things being equal, would score 50 points better on the biochemistry subtest than someone from the school we chose for the baseline. We also

STUDENT PERFORMANCE ANALYSIS

- NATIONAL BOARD OF MEDICAL EXAMINERS EXAM SCORES
- 10 SCHOOLS, 1975 AND 1976 CLASSES (1900 MAJORITY, 200 MINORITY STUDENTS
- STUDENT ATTRIBUTES PREDICT PERFORMANCE
- SEPARATE MINORITY AND MAJORITY ANALYSES

looked at data on when students took the test because, as most of you may know, the National Board Exams are given at different times in different places. And, finally, we performed separate minority and majority analyses.

Now before talking about the majority and minority performance prediction equations, let me be clear that there was no expectation that the two groups would perform equally because the two groups are not the same when they come into school. (Slide 11) This table just shows means for the minority and majority groups for the Part II exam takers. You see that the average minority student has a lower GPA, lower MCAT, and so on. So, not surprisingly, that person, as we would expect, performed less well on the Part II exam. Incidentally, the 498 for majorities is almost exactly at the national mean; it is a standardized test score with mean 500 and 100 standard deviation. You see that the minority mean is lower. The purpose of the analysis was not to determine something that people already knew, i.e., that if people come in less well prepared, they will do less well. Rather, the question was, "What factors predict performance and how do these factors differ for the two groups."

Let me summarize this without giving you a lot of numbers. (Slide 12) For the majority students, traditional factors predict performance. The most significant predictor is science GPA, followed by science MCAT, followed by selectivity of an undergraduate school, followed in very close proximity by general information MCAT, verbal MCAT, and quantitative MCAT.

SOME AVERAGE STUDENT CHARACTERISTICS

(PART II EXAM TAKERS)

	MEANS	
	MAJORITY	MINORITY
UNDERGRADUATE GPA	3.39	3.11
VERBAL MCAT	580	515
GENERAL INFORMATION MCAT	571	496
QUANTITATIVE MCAT	634	573
SCIENCE MCAT	586	526
PART II SCORE	498	424

PREDICTING STUDENT PERFORMANCE

- MAJORITY: ACHIEVEMENT MEASURES
- MINORITY: ACHIEVEMENT AND APTITUDE MEASURES
- MINORITY: ADMISSION AND PERFORMANCE DIFFER

Now the picture for the minority students is really quite different. The verbal MCAT is better than the science MCAT or science GPA on five out of six of the tests. That is, the score of a minority student's verbal MCAT was a better overall predictor than the commonly used admissions criteria. Recall the earlier slide showing admissions equations with the multicolor coding; verbal MCAT was rarely used to admit minority applicants. Nonetheless, it is a very significant predictor of performance. The quantitative MCAT was also a good predictor of minority performance. It is particularly interesting that verbal and quantitative MCATS do generally quite well in predicting performance on the clinical part of the test. After you control for the scores of the minority student, the selectivity of the undergraduate school attended is not a significant predictor of performance.

To me, the most interesting aspect of this analysis is the comparison between factors that predict admission and those that predict performance. If you look back at the admissions criteria for majority applicants, they are very similar to those that predict performance of majority medical students. That is not the case for minority students. Indeed, the verbal and quantitative MCAT are in general better than, but are at least equivalent to, the standard science scores as predictors of minority student scores.

Now what about minority and majority comparisons? (Slide 13) If we take two individuals who enter medical school with the same preadmission academic characteristics, a hypothetical minority and majority student, we would find on average the majority student scores about 40 points higher than the minority student. However, before you make very much of that, look at it in terms of the significance of that spread. The odds would be only about three to two that your majority student would outscore the equivalent minority student.

How can we explain these results? Quite frankly they were surprising to us. Although there are no explanations that have been tested empirically, we offer some hypotheses. (Slide 14)

The first explanation relates to our analysis of minority and majority admissions. We find that admissions committees did not, at least by this analysis, use quantitative preadmission characteristics as well in selecting minority students who would perform well as they did in selecting majority students. I would also guess that admissions committees probably are not as good at judging qualitative factors that predict performance for minority students as they are for majority students. They simply haven't had as much experience.

Another possible explanation is inflated GPA. If, as some have hypothesized, the undergraduate GPA of minority students had been inflated, then the model (using inflated GPA) would predict minority students to perform better than their "true" GPA would lead one to expect. I think that this may be a partial explanation for our results but it certainly cannot fully explain

MINORITY/MAJORITY COMPARISONS

- ON AVERAGE, MAJORITY STUDENT SCORES 40 POINTS HIGHER GIVEN EQUAL PREADMISSION CHARACTERISTICS

- INDIVIDUALLY, MAJORITY STUDENT OUTSCORES MINORITY WITH EQUAL PREADMISSION CHARACTERISTICS 2 OUT OF 3 TIMES

POSSIBLE EXPLANATIONS OF MINORITY/MAJORITY DIFFERENCES

- ADMISSION DECISIONS
- INFLATED MINORITY GPA
- PRESSURE TO KEEP MINORITIES
- DIFFERING CULTURAL EFFECTS

the difference observed.

What about cultural bias on the test? That is possible. The National Board test would have to be more culturally biased than the MCAT for this to serve as an explanation. That, to me at least, does not seem particularly plausible because the MCAT is much more a broad achievement test. The old General Information test, for example, clearly had in it some cultural questions.

Those are our hypotheses, and I think it is worth examining the question further.

The last research that I will describe goes far down stream in the medical education continuum and deals with the American Board of Internal Medicine (ABIM) certifying exam. (Slide 15) Again, we have nine of our ten schools but for different reasons. One of the ten schools started too late to have graduates in our cohort, the classes are 1955, 1960, 1965, and 1969. The data we use to predict performance on the ABIM test include the standard kinds of preadmissions data, MCATs, GPA, and so on, which were obtained from the records of the nine schools. In our analysis, we also use medical school performance, i.e., the National Board exams, rank in class, and postgraduate training, i.e., internship, residency in the major teaching hospitals. The exam results are reported as standardized scores in the same basic format as the MCAT and National Board scores.

ANALYSIS OF ABIM CERTIFYING EXAM

SUBJECTS

438 Medical School Graduates taking ABIM exam
- Nine schools
- Classes of 1955, 1960, 1965, 1969

DATA

Preadmission data (MCATs, GPA, etc.)
Medical School Performance (National Board scores, Rank, etc.)
Postgraduate training (internship, residency in major teaching hospital)

American Board of Internal Medicine certifying exam results

As an aside, ABIM used to give an exam that involved an oral portion. In the Philadelphia office of the ABIM, there is a little tombstone that says, "Here lies the oral exam, victim of a low r (low correlation)." Very simply, they found very low interrater reliability among those who evaluated oral exam results. That is, the two doctors or the three doctors who gave the test did not agree sufficiently in their assessments of the applicant for the ABIM to defend the value of the oral test. So they gave up on the oral, qualitative assessment of patient management, and now rely wholly on a standardized test.

Results of our analysis are fairly straightforward. (Slide 16) If you use rank in medical school class, age, and school differences --that is a variable to account for the difference between Duke and school X--you can explain 30% of the variance in scores on the American Board of Internal Medicine certifying exam. The effects of the MCAT, undergraduate GPA, and National Boards are still significant after you control for rank and class, age and school.

Let me emphasize that the elapsed time between entering medical school and taking this exam ranges from a minimum of seven years to a maximum, I think, of ten or fifteen years. So we're talking about predictive value of those standardized tests really far down the pike in medical education. The certifying exam is sort of the final hurdle in the medical education process.

RESULTS OF ANALYSIS

- Rank in medical school class, age, and school differences account for 30 percent of variance
- Effects of MCAT, Undergrad GPA, National Boards are still significant after controlling for rank, age, school
- The longer applicants wait to take exam, the worse their performance
- Relatively simple model accounts for 55 to 65 percent of explainable variance
- Relationship between board certification and favorable PSRO review suggests knowledge translates to actual practice

Now an interesting thing, at least to us, was that the longer you wait to take the exam after you could take it, the worse you do. You can think of that as a selection bias. The applicant may think, "I know I'm not going to do well on it, I have anxiety, I'll wait." The alternative hypothesis is forgetting.

Now consider our results not simply in terms of the amount of variance the model predicts but also in terms of the error in the test. This exam--indeed every exam--does not test those things that it tries to test perfectly. We find that we can predict 55 to 65% of the explainable variance with a fairly simple model.

Finally, what does this mean, or who cares about the certifying exam? Obviously, people who are trying to get certified care about it. What is the evidence that this makes a difference in the delivery of patient care? Well, the evidence is limited. However, studies of Professional Standard Review Organizations (PSRO) have found that, after you control for age and other characteristics, board certified physicians provide better care. That is, assessments of quality of care by peers show that people who are board certified deliver better care as judged by their peers.

Let me give you the general conclusions from these studies.

- o First, analysis of the admissions process can really help a school evaluate its criteria. For example, some of the schools were weighting the state of residence more heavily than they realized.

- o Second, analysis of the sort that we did on state of residence can help the student evaluate his or her chances of admission.
- o Third, analysis of performance data can lead to better student selection in affirmative action programs. Schools can use such analyses to select, from those applicants who apply, students who will do better and have less trouble in medical school. Now I would not suggest for a minute that admission committees should be concerned only about the difficulty someone has in getting through medical school. However, analysis of this sort can at least let you look at that and predict it better. In the mid 1970s medical schools were using different criteria for minority admissions than those that predict minority performance.
- o Fourth, I think the analysis of the certifying exam data suggests the MCAT and the National Board test are at least valid predictors of performance on a down-the-pike exam of clinical performance. I should be clear. That is still a paper and pencil test but it deals with patient management problems and other measures of clinical knowledge.
- o Fifth, multi-school collaborative studies are essential. It is next to impossible for a single school to look at a single class and do an analysis that will tease out all the factors that predict performance. You simply have too few data, and that is particularly true for minority

applicants. If you have a class of 10, 20, or 30, it is simply not possible, statistically, to do an analysis that will let you know, other than by chance, what predicts performance. Medical schools really need to pool their data in order to understand more about minority performance, particularly. The new MCAT and the re-certifying exams are very interesting targets for further analysis.

- o Finally, I would like to restate my strong plea that we do more analysis of qualitative data. The kinds of analysis we do are relatively easy, because they involve numbers. There are a whole lot of other things that are important in the admissions process and in performance that really have not been quantified sufficiently to permit sound analysis.